63-2-6

D1-82-0234

"Also available from the author"

BOEING SCIENTIFIC RESEARCH LABORATORIES

Expressing the Normal Distribution with Covariance Matrix A+B in Terms of One with Covariance Matrix A.

George Marsaglia

Mathematics Research

February 1963

D1-82-0234

EXPRESSING THE NORMAL DISTRIBUTION WITH COVARIANCE MATRIX $A + B$ IN TERMS OF ONE WITH COVARIANCE MATRIX A.

by

George Marsaglia

Mathematical Note No. 287

Mathematics Research Laboratory

BOEING SCIENTIFIC RESEARCH LABORATORIES

February 1963

1. Introduction

References [1] - [6] list a few of many papers concerned with evaluating the multivariate normal integral. I have singled out these because part of each of them is devoted to the problem of obtaining a particular type of reduction formula for the normal integral - writing it in terms of more tractable normal integrals. In spite of the apparent variety of results and methods of these papers, they all give formulas which are actually simple applications of the convolution formula (1) given below. The purpose of this note, then, is to point out that a number of reduction formulas, derived with various degrees of difficulty in the literature, are special cases of formula (1), and to point out that many more reduction formulas easily follow from (1). The ease with which formulas may be derived suggests that most of them are mere curiosities, and of little value for numerical work, but the general applicability of the formula makes it worth having as a tool for attacking particular normal distributions with given covariance matrices. In addition, from a variety of representations of a normal integral as an expectation, one would hope to be able to choose one with small variance for Monte Carlo estimates.

2. The Formula

If x and y are independent random variables with distributions F and G, then the distribution of $x + y$ may be expressed in this form:

$$P[x+y < a] = \int P[x < a-y|y]dG = \mathcal{E}\{F(a-y)\}$$

where $\mathcal{E}$ is expectation. This same form of the convolution formula applies

if $\xi = (x_1,\dots,x_n)$ and $\eta = (y_1,\dots,y_n)$ are independent $1\times n$ vectors – if $\alpha = (a_1,\dots,a_n)$ and $F(\alpha) = P[\xi < \alpha] = P[x_1 < a_1,\dots,x_n < a_n]$, then

$$P[\xi+\eta < \alpha] = \mathcal{E}\{F(\alpha-\eta)\}.$$

Now let $\mathrm{cov}(\xi) = A$ mean that A is the covariance matrix of the vector ξ. If ξ and η are independent $1\times n$ vectors of zero mean normal random variables with $\mathrm{cov}(\xi) = A$ and $\mathrm{cov}(\eta) = B$ then $\mathrm{cov}(\xi+\eta) = A+B$, and if we let $F(A,\alpha) = P[\xi < \alpha]$, then

$$F[A+B,\alpha] = P[\xi+\eta < \alpha] = \mathcal{E}\{F(A,\alpha-\eta)\}$$

This formula may be used to give most of the reduction formulas reported in the literature, so we state it precisely.

<u>Let</u> $F(A,\alpha)$ <u>be</u> <u>the n-dimensional</u> <u>normal</u> <u>distribution</u>,

$$F(A,\alpha) = P[\xi < \alpha] = P[x_1 < a_1,\dots,x_n < a_n],$$

<u>where</u> $\xi = (x_1,\dots,x_n)$ <u>and</u> <u>the</u> x's <u>are</u> <u>jointly</u> <u>normal</u> <u>with</u> <u>zero</u> <u>means</u> <u>and</u> <u>covariance</u> <u>matrix</u> A. <u>Then</u>

$$F(A+B,\alpha) = \mathcal{E}\{F(A,\alpha-\eta)\} \tag{1}$$

<u>where</u> $\eta = (y_1,\dots,y_n)$ <u>is</u> <u>a</u> <u>zero</u> <u>mean</u> <u>normal</u> <u>vector</u> <u>with</u> $\mathrm{cov}(\eta) = B$.

3. <u>Some</u> <u>Notation</u>

Let Φ be the standard normal integral, so that

$$F(I,\alpha) = \Phi(a_1)\Phi(a_2)\ \dots\ \Phi(a_n),$$

and let H be the Heaviside function, or the characteristic function of

the positive orthant - if $\beta = (b_1,\ldots,b_m)$ then $H(\beta) = 1$ if $\beta > 0$ or $H(\beta) = 0$ otherwise. For a partitioned covariance matrix with blocks of zeros, $\binom{R\ 0}{0\ 0}$, and a partitioned vector (α,β), we have

$$(2) \qquad F[\begin{pmatrix} R & 0 \\ 0 & 0 \end{pmatrix}, (\alpha,\beta)] = F(R,\alpha)H(\beta).$$

Also note that if ν is the vector of all 1's, $\nu = (1,\ldots,1)$, then

$$F(\nu'\nu,\alpha) = \Phi[\min(a_1,\ldots,a_n)],$$

since if y has distribution Φ, then $y\nu = (y,y,\ldots,y)$ has covariance matrix $\nu'\nu$. In general, if $\text{cov}(\eta) = B$, then $\text{cov}(\eta M) = M'BM$.

4. Applications

We give some examples of the use of (1). Suppose that A is non-singular. Multiplying by a constant if necessary, we may assume that the smallest characteristic root of A is 1. Then A has a representation $A = I+T'T$ where T is kxn of rank $k < n$. The multiplicity of the root 1 is n-k. Then

$$F(A,\alpha) = F(I+T'T,\alpha) = \mathcal{E}\{F(I,\alpha-\zeta T)\}$$

where $\zeta = (z_1,\ldots,z_k)$ has zero means and covariance I. Hence if $\tau_1',\tau_2',\ldots,\tau_n'$ are the columns of T,

$$(3) \qquad F(A,\alpha) = F(I+T'T,\alpha) = \mathcal{E}\{\Phi(a_1-\zeta\tau_1')\Phi(a_2-\zeta\tau_2')\ldots\Phi(a_n-\zeta\tau_n')\}$$

A particularly easy example of (3) is when $k = 1$, so that $A = I+\tau'\tau$, with $\tau = (t_1,\ldots,t_n)$. Then

$$(4) \qquad F(I+\tau'\tau,\alpha) = \mathcal{E}\{\Phi(a_1+t_1z)\Phi(a_2+t_2z)\ldots\Phi(a_n+t_nz)\}$$

with z having distribution Φ. Formula (3) includes results of Das [2], Moran [5], and Stuart [6] ; Ihm [3], went to some length to establish a version of (4).

Again multiplying by a constant if necessary, we may assume that $A-\nu'\nu$ is non-negative definite, where once again $\nu = (1,1,\ldots,1)$. Then

$$F(A,\alpha) = F(A-\nu'\nu+\nu'\nu,\alpha) = \mathcal{E}\{F(\nu'\nu,\alpha-\eta)\}, \ \mathrm{cov}(\eta) = A-\nu'\nu.$$

Hence

$$F(A,\alpha) = \mathcal{E}\{\Phi[\min(a_1-y_1,a_2-y_2,\ldots,a_n-y_n)]\}$$

where $y_1,\ldots,y_n$ are normal, zero means, covariance $A-\nu'\nu$. More generally, if we write $A = A-\gamma'\gamma+\gamma'\gamma$ in such a way that $A-\gamma'\gamma$ is non-negative definite, with $\gamma = (c_1,\ldots,c_n)$, then

$$F(A,\alpha) = \mathcal{E}\{F(\gamma'\gamma,\alpha-\eta)\}, \quad \mathrm{cov}(\eta) = A-\gamma'\gamma.$$

It is easy to deal with $F(\gamma'\gamma, \ )$ for example, if all the c's are positive,

$$F(A,\alpha) = \mathcal{E}\{F(\gamma'\gamma,\alpha-\eta)\} = \mathcal{E}\{\Phi[\min(c_1^{-1}a_1-c_1^{-1}y_1,\ldots,c_n^{-1}a_n-c_n^{-1}y_n)]\}$$

where $y_1,\ldots,y_n$ are normal, zero means, covariance $A-\gamma'\gamma$. We may always choose γ so that $A-\gamma'\gamma$ has rank less than that of A.

Another type of reduction formula involves integrals over half-infinite intervals. To develop these, let the composite zero mean normal vector (ξ,η) have covariance $\begin{pmatrix} A & B' \\ B & C \end{pmatrix}$, with $\operatorname{cov}(\xi) = A$ and $\operatorname{cov}(\eta) = C$. Assume C^{-1} exists. Now write

$$P[(\xi,\eta) < (\alpha,\beta)] = F[\begin{pmatrix} A & B' \\ B & C \end{pmatrix}, (\alpha,\beta)] = F[\begin{pmatrix} A-B'C^{-1}B & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} B'C^{-1}B & B' \\ B & C \end{pmatrix}, (\alpha,\beta)],$$

then use (1) to get

$$F[\begin{pmatrix} A & B' \\ B & C \end{pmatrix}, (\alpha,\beta)] = \mathcal{E}\{F[\begin{pmatrix} A-B'C^{-1}B & 0 \\ 0 & 0 \end{pmatrix}, (\alpha-\eta C^{-1}B, \beta-\eta)]\},$$

since the vector $(\eta C^{-1}B, \eta) = \eta(C^{-1}B, I)$ has covariance $\begin{pmatrix} B'C^{-1}B & B' \\ B & C \end{pmatrix}$ if η has covariance C. Using (2) to remove the blocks of zeros, we have our result:

$$(5)\quad F[\begin{pmatrix} A & B' \\ B & C \end{pmatrix}, (\alpha,\beta)] = \mathcal{E}\{F(A-B'C^{-1}B, \alpha-\eta C^{-1}B) H(\beta-\eta)\}, \quad \operatorname{cov}(\eta) = C.$$

Special cases of (5) are reported in [1] and [4]. Note also that (5) may be used to establish that the conditional mean and covariance of ξ, given η, are $\eta C^{-1}B$ and $A-B'C^{-1}B$. We can, in fact, give a more general form of (5) which does not require that C^{-1} exist and which will give the most general formulas for conditional mean and covariance. Let C^+ be the psuedo inverse of C, that is, if $C = T'T$ with T $r \times m$ of rank r, then $C^+ = T'(TT')^{-2}T$, alternatively, $C^+ = RR'$ where $R = T'(TT')^{-1}$ is the principal right inverse of T.

Using (1), we have

$$F[\begin{pmatrix} A & B' \\ B & C \end{pmatrix}, (\alpha,\beta)] = \mathcal{E}\{F[\begin{pmatrix} A-B'C^+B & 0 \\ 0 & 0 \end{pmatrix}, (\alpha-\eta C^+B, \beta-\eta)]\}$$

since the vector $(\eta C^+B, \eta) = \eta(C^+B, I)$ has covariance $\begin{pmatrix} B'C^+B & B' \\ B & C \end{pmatrix}$.

(A direct calculation shows that the covariance is $\begin{pmatrix} B'C^+B & B'E \\ EB & C \end{pmatrix}$, where $E = T'(TT')^{-1}T$ is the projector of the row space of C. But $CE = C$, and hence for any matrix H whose rows are in the row space of C, $HE = H$. In particular, $B'E = B'$, since the rows of B' lie in the row space of C. To prove this last contention, note that any covariance matrix may be assumed to have the form $\begin{pmatrix} S' \\ V' \end{pmatrix} (SV) = \begin{pmatrix} S'S & S'V \\ V'S & V'V \end{pmatrix}$, and the rows of $S'V$ lie in the row space of $V'V$, which is the row space of V.)

Using (2) to remove blocks of zeros, we have a more general form of (5):

(6) $$F[\begin{pmatrix} A & B' \\ B & C \end{pmatrix},(\alpha,\beta)] = \mathcal{E}\{F(A-B'C^+B,\alpha-\eta C^+B)H(\beta-\eta)\}, \quad \mathrm{cov}(\eta) = C.$$

From (6) follows the general formula that the conditional mean and covariance of ξ, given η, are ηC^+B and $A-B'C^+B$.

REFERENCES

[1] Curnow, R. N. and Dunnett, C. W., "The numerical evaluation of certain multivariate normal integrals," Ann. Math. Stat. 33, 571 - 579.

[2] Das, S. C. (1956) "The numerical evaluation of a class of integrals", II, Proc. Camb. Philos. Soc. 52, 442 - 448.

[3] Ihm, P. (1959) "Numerical evaluation of certain multivariate normal integrals", Sankhyā 21, 363 - 366.

[4] John, S. "On the evaluation of the probability integral of a multivariate normal distribution," Sankhyā 21, 367 - 370.

[5] Moran, P.A.P. (1956) "The numerical evaluation of a class of integrals". Proc. Camb. Philos, Soc 52 ,230 - 233.

[6] Stuart, Alan (1958) "Equally correlated variates and the multinormal integral," J. Roy. Statist. Soc., Ser. B 20 ,373 - 378.